Challenges and Proposals on Using LLM Gen AI for Financial Risk Analysis *

*Note: Sub-titles are not captured for https://ieeexplore.ieee.org and should not be used

1st Shivgan Joshi
QcFinance IN
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—In this paper we have proposed and demonstrated a protype on how to use GenAI in modeling Financial Risk Analysis.**

Model development, validation and approval is an art and science. Mathematical models without human oversight can lead to failure. In this research we will look at how to use the current development in GenAi LLM to risk modeling.

This document is a model and instructions for LaTeX. This and the IEEEtran.cls file define the components of your paper [title, text, heads, etc.]. *CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract.

***Index Terms*—Gen AI, Financial Risk, style, styling, insert.**

I. Introduction

A lot of information about risk modeling is already in public domains and ChatGPT kind of tools already have them in their models. There are two ways to develop models, either use Open AI and use that in the Org or train your own models. We will be looking at research to find out what people are suggesting.

This document is a model and instructions for LaTeX. Please observe the conference page limits. For more information about how to become an IEEE Conference author or how to write your paper, please visit IEEE Conference Author Center website: https://conferences.ieeeauthorcenter.ieee.org/.

A. Maintaining the Integrity of the Specifications

The IEEEtran class file is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations.

Identify applicable funding agency here. If none, delete this.

II. Literature Review

Teixeira, Marar, Yazdanpanah, *et al.* [1] discusses credit risk models and their integration with large language models to enhance report generation. They uses the data collected mostly collected from conversing with the potential applicant and convert it into data points and then calculate factors like Probability of Default (PD). In other words, data collected while talking to the customer is then used on the fly on trained models to calculate loan eligibility. This is more the on lead generation side.

Sanz-Guerrero and Arroyo [2] extensively uses bert llm huggingface to on Peer to Peer lending on application data.

Khoja [3] uses LLM for Bond Valuation. Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections II-E to II-I below for more information on proofreading, spelling and grammar.

Babaei and Giudici [4] has discussed new ideas.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads—LaTeX will do that for you.

A. Synthesis Matrix on Topic A

Table II provides a synthesis of the key literature reviewed, organized by objective, methodology, and findings/gaps.

B. Synthesis Matrix on Topic B

Table II summarizes key literature, focusing on objectives, methodologies, and findings/gaps.

TABLE I
SYNTHESIS MATRIX FOR LITERATURE REVIEW TOPIC A

Reference	Objective/Focus	Methodology	Key Findings/Gaps
ref1	Investigate the relationship between X and Y	Statistical regression analysis on dataset Z	Identified significant correlation; lacks causal analysis
ref2	Develop a framework for A in context B	Simulation and modeling	Improved system efficiency by 15%; limited scalability
ref3	Analyze impact of C on D	Mixed methods: Surveys and case studies	Found diverse regional trends; requires larger sample size
ref4	Evaluate effectiveness of E compared to F	Experimental study with control group	E outperformed F in accuracy but had higher costs

TABLE II
SYNTHESIS MATRIX FOR LITERATURE REVIEW TOPIC B

Reference	Objective/Focus	Methodology	Key Findings/Gaps
ref1	Investigate the relationship between X and Y	Statistical regression analysis on dataset Z	Identified significant correlation; lacks causal analysis
ref2	Develop a framework for A in context B	Simulation and modeling	Improved system efficiency by 15%; limited scalability
ref3	Analyze impact of C on D	Mixed methods: Surveys and case studies	Found diverse regional trends; requires larger sample size
ref4	Evaluate effectiveness of E compared to F	Experimental study with control group	E outperformed F in accuracy but had higher costs

C. Model Comparison

Table III provides a comparison of various research models based on objectives, methods, datasets, and performance metrics.

TABLE III
COMPARISON OF RESEARCH MODELS

Model	Objective	Methodology	Dataset	Performance
model1	Predict future trends in X	Neural networks with feature selection	Dataset A (1M samples)	Accuracy: 92%, Precision: 89%
model2	Classify instances of Y in real-time	Random forest classifier	Dataset B (500K samples)	F1-Score: 85%, Recall: 87%
model3	Optimize Z under constraints	Linear programming with heuristics	Dataset C (100K samples)	Computational efficiency improved by 30%
model4	Detect anomalies in W	Autoencoder with anomaly score thresholding	Dataset D (200K samples)	AUC: 0.94, Sensitivity: 91%

D. Yes/No Chart and Heat Map for Literature Review Topic X

Table IV summarizes the features and criteria addressed by the reviewed studies.

TABLE IV
YES/NO CHART FOR LITERATURE REVIEW

Criteria/Feature	ref1	ref2	ref3	ref4
Addresses Problem X	Yes	Yes	No	Yes
Uses Dataset Y	No	Yes	Yes	No
Employs Method Z	Yes	No	No	Yes
Reports Performance Metrics	Yes	Yes	Yes	No
Proposes a New Framework	No	Yes	No	No

TABLE V
HEATMAP FOR LITERATURE REVIEW

Criteria/Feature	ref1	ref2	ref3	ref4
Addresses Problem X	High	Medium	Low	Very Low
Uses Dataset Y	Low	High	Medium	Low
Employs Method Z	Medium	Low	Very Low	High
Reports Performance Metrics	High	High	Medium	Low
Proposes a New Framework	Very Low	Medium	Low	Very Low

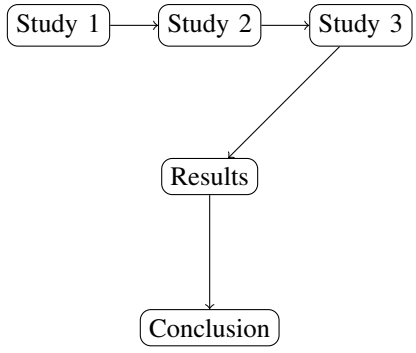
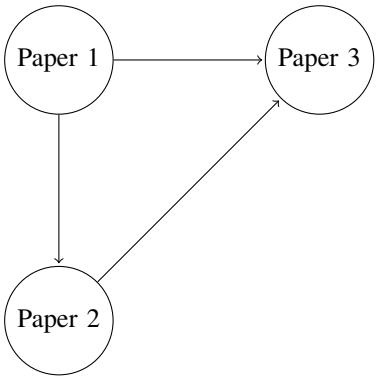
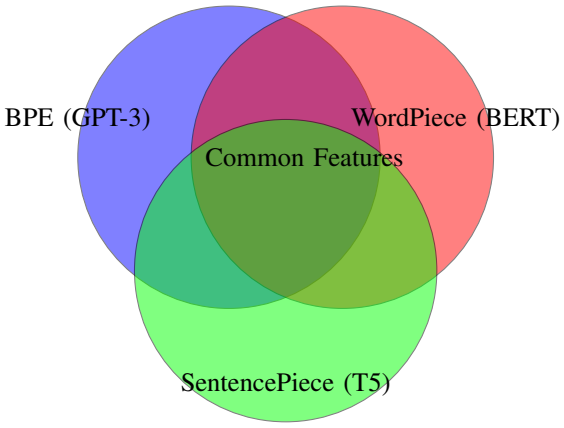


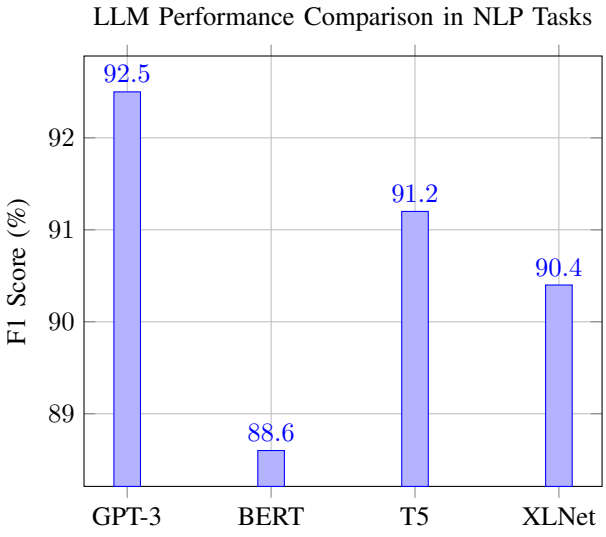
Fig. 1. Flow Diagram of Studies and Results



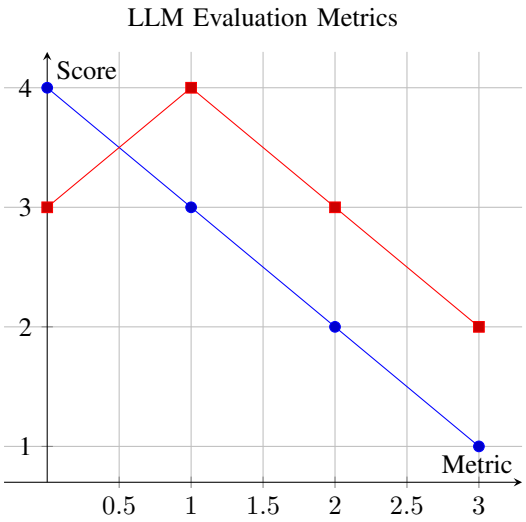


fdfdfdsfas

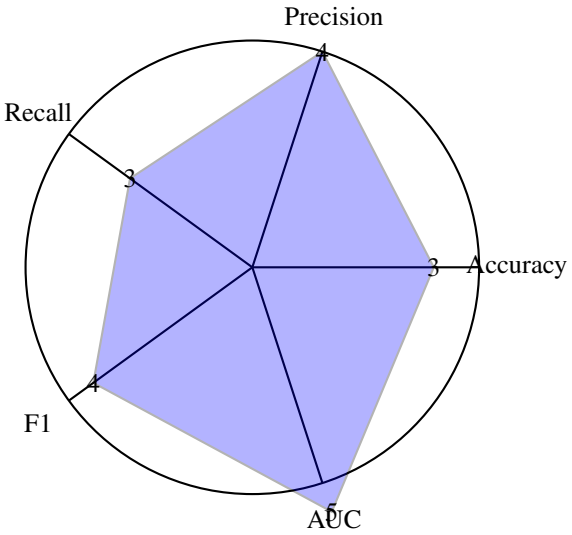
twat



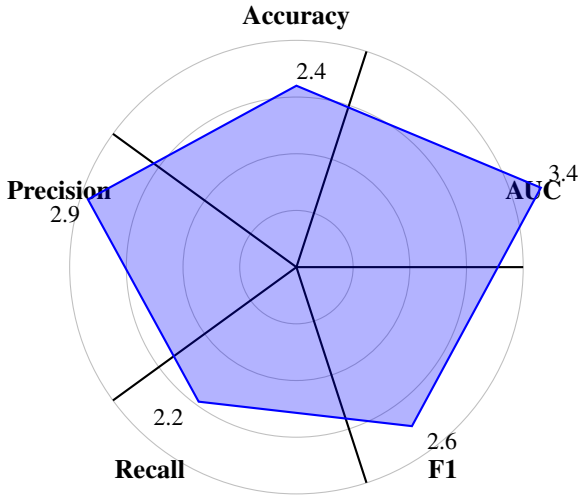
fdg gfsgsfg gs ggs



qewq eq eqe qeq



ffdsfsfsv



Radar Chart of LLM Evaluation Metrics

gsgdsgsdg

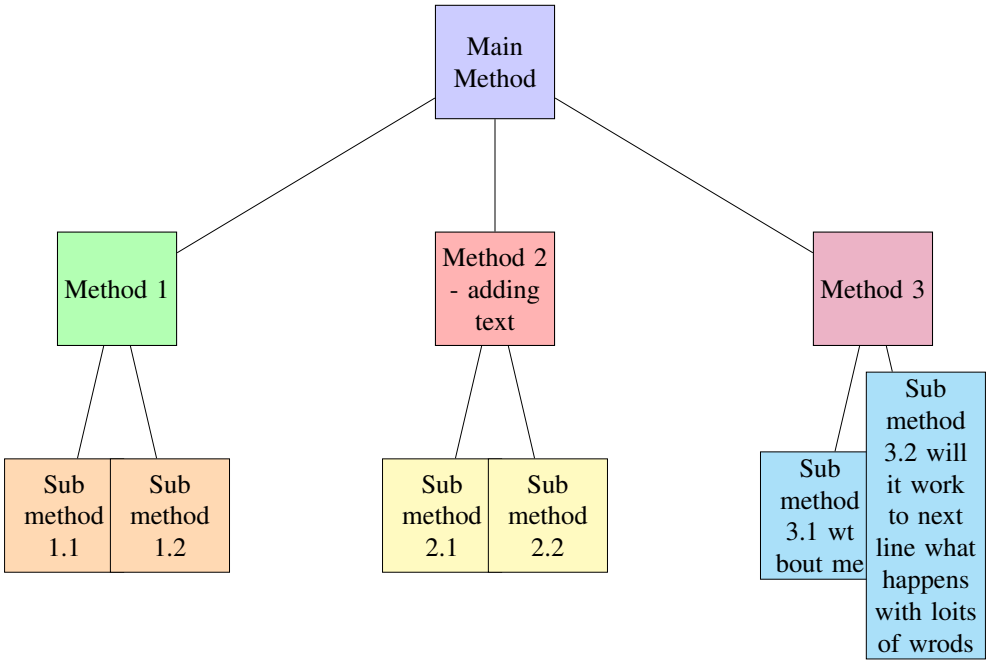


Fig. 2. Methods Tree Diagram in Literature

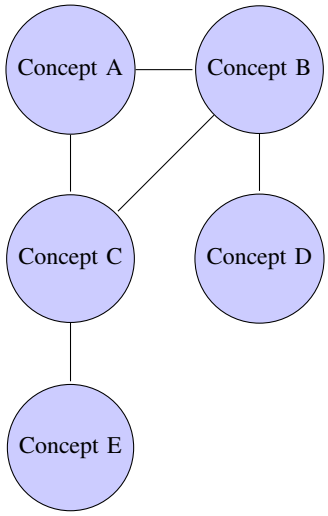


Fig. 3. Connected Chart of Methods and Concepts

small heatmap
affaf ar ara ra rar a

E. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, ac, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

F. Units

- Use either SI (MKS) or CGS as primary units. (SI units are encouraged.) English units may be used as secondary units (in parentheses). An exception would be the use of English units as identifiers in trade, such as “3.5-inch disk drive”.
- Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity that you use in an equation.
- Do not mix complete spellings and abbreviations of units: “Wb/m^2” or “webers per square meter”, not “webers/m^2”. Spell out units when they appear in text: “. . . a few henries”, not “. . . a few H”.
- Use a zero before decimal points: “0.25”, not “.25”. Use

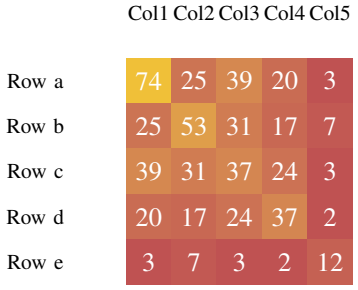


Fig. 4. 5x5 Heatmap Example

"cm^3", not "cc".)

G. Equations

Number equations consecutively. To make your equations more compact, you may use the solidus (/), the exp function, or appropriate exponents. Italicize Roman symbols for quantities and variables, but not Greek symbols. Use a long dash rather than a hyphen for a minus sign. Punctuate equations with commas or periods when they are part of a sentence, as in:

$$a + b = \gamma \tag{1}$$

Be sure that the symbols in your equation have been defined before or immediately following the equation. Use "(1)", not "Eq. (1)" or "equation (1)", except at the beginning of a sentence: "Equation (1) is . . ."

H. LaTeX-Specific Advice

Please use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). That will make it possible to combine sections, add equations, or change the order of figures or citations without having to go through the file line by line.

Please don't use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Please note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed. If you forget that, you might write an article in which the equation numbers skip from (17) to (20), causing the copy editors to wonder if you've discovered a new method of counting.

LaTeX can't read your mind. If you assign the same label to a subsubsection and a table, you might find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` (there won't be any anyway) and it might stop a wanted equation number in the surrounding equation.

I. Figures and Tables

a) Positioning Figures and Tables: Place figures and tables at the top and bottom of columns. Avoid placing them in the middle of columns. Large figures and tables may span across both columns. Figure captions should be below the figures; table heads should appear above the tables. Insert figures and tables after they are cited in the text. Use the abbreviation "Fig. 5", even at the beginning of a sentence.

TABLE VI
TABLE TYPE STYLES

Table Head	**Table Column Head**		
	Table column subhead	***Subhead***	***Subhead***
copy	More table copy[a]		

[a]Sample of a Table footnote.

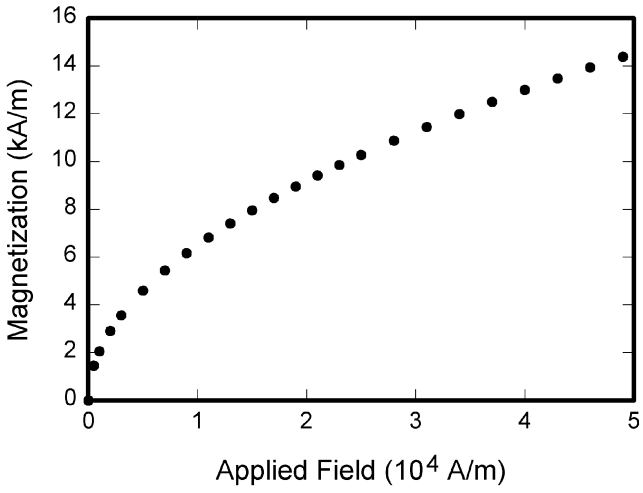


Fig. 5. Example of a figure caption.

Figure Labels: Use 8 point Times New Roman for Figure labels. Use words rather than symbols or abbreviations when writing Figure axis labels to avoid confusing the reader. As an example, write the quantity "Magnetization", or "Magnetization, M", not just "M". If including units in the label, present them within parentheses. Do not label axes only with units. In the example, write "Magnetization (A/m)" or "Magnetization {A[m(1)]}", not just "A/m". Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)", not "Temperature/K".

III. Proposed Methodology

We propose to integrate like code assist, model assist to the modeling framework to ease developing regulating and testing models.

IV. Conclusion

We have proposed on the optimal way to integrate current Gen Ai in risk modeling.

Acknowledgment

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks . . .”. Instead, try “R. B. G. thanks. . .”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

References

[1] A. C. Teixeira, V. Marar, H. Yazdanpanah, A. Pezente, and M. Ghassemi, “Enhancing credit risk reports generation using LLMs: An integration of bayesian networks and labeled guide prompting,” in *Proceedings of the Fourth ACM International Conference on AI in Finance*, ser. ICAIF ’23, New York, NY, USA: Association for Computing Machinery, Nov. 25, 2023, pp. 340–348, ISBN: 979-8-4007-0240-2. DOI: 10.1145/3604237.3626902. [Online]. Available: https://dl.acm.org/doi/10.1145/3604237.3626902 (visited on 12/19/2024).

[2] M. Sanz-Guerrero and J. Arroyo, *Credit risk meets large language models: Building a risk indicator from loan descriptions in p2p lending*, Aug. 5, 2024. DOI: 10.48550/arXiv.2401.16458. arXiv: 2401.16458[q-fin]. [Online]. Available: http://arxiv.org/abs/2401.16458 (visited on 12/19/2024).

[3] M. Khoja, *AI and bond values: How large language models predict default signals*, Rochester, NY, Sep. 20, 2024. DOI: 10.2139/ssrn.4965227. [Online]. Available: https://papers.ssrn.com/abstract=4965227 (visited on 12/19/2024).

[4] G. Babaei and P. Giudici, “GPT classifications, with application to credit lending,” *Machine Learning with Applications*, vol. 16, p. 100 534, Jun. 1, 2024, ISSN: 2666-8270. DOI: 10.1016/j.mlwa.2024.100534. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666827024000100 (visited on 12/19/2024).